Engineers develop framework to predict types of sounds likely to be heard at certain locations

November 22 2023, by Shawn Ballard

A soundscape map (left) for the text prompt, "This is a sound of sea waves," and the region's corresponding overhead image. Green indicates areas where the sound is more probable; white indicates less probable. Credit: Jacobs lab

Imagine yourself on a beautiful beach. You're likely visualizing sand and sea but also hearing a symphony of wind gusting, waves crashing and gulls cawing. In this scene—as well as in urban settings with neighbors talking, dogs barking and traffic whooshing—sounds are critical

components of the overall feel of a place.

Indeed, sound is one of the fundamental senses that helps humans understand their environments, and environmental sound conditions have been shown to have a strong correlation with a person's mental and physical health. Reliable methods for understanding the soundscape of a given geographic area are therefore valuable for applications ranging from collective policymaking around urban planning and noise management to individual decisions about where to buy a home or establish a business.

Nathan Jacobs, a professor of computer science and engineering, along with graduate students Subash Khanal, Srikumar Sastry and Aayush Dhakal, all studying computer science and engineering, at the McKelvey School of Engineering at Washington University in St. Louis, developed Geography-Aware Contrastive Language Audio Pre-training (GeoCLAP), a novel framework for soundscape mapping that can be applied anywhere in the world.

They presented their work on Nov. 22 at the British Machine Vision Conference in Aberdeen, United Kingdom. The paper is also posted to the *arXiv* preprint server.

The team's key innovation comes from their use of three different modalities, or types of data, in their framework, which incorporates geotagged audio, textual description and overhead images. Unlike previous methods for soundscape mapping that focused on only two modalities, GeoCLAP's richer understanding allows users to create probable soundscapes from either textual or audio queries for any geographic location.

"We've developed a simple and scalable way of creating a soundscape map for any geographic area," Jacobs said. "Our approach overcomes

the limitations of previous soundscape mapping methods that were rule-based, often missing important sound sources, or relied on direct human observations, which are difficult to obtain in sufficient quantities away from popular tourist destinations.

"By leveraging the intrinsic relationship between sound and localized visual cues, our multimodal tool and freely available overhead imagery makes it possible for us to create soundscape maps for any area in the world."

Provided by Washington University in St. Louis